

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: MOORE, PAUL A.
RUBEN, STEVEN M.
EBNER, REINHARD
- (ii) TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEIN
- (iii) NUMBER OF SEQUENCES: 16
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: HUMAN GENOME SCIENCES, INC.
(B) STREET: 9410 KEY WEST AVENUE
(C) CITY: ROCKVILLE
(D) STATE: MD
(E) COUNTRY: USA
(F) ZIP: 20850
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: BROOKES, ANDERS A.
(B) REGISTRATION NUMBER: 36,373
(C) REFERENCE/DOCKET NUMBER: PF378
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (301) 309-8504
(B) TELEFAX: (301) 309-8512

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2329 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 124..913
- (ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 124..184
- (ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 187..913

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TTACCAGAAC	AGCATAACAA	GGGCAGGTCT	GACTGCAAGC	TGGGACTGGG	AGGCAGAGCC	60
GCCGCCAAGG	GGGCCTCGGT	TAAACACTGG	TCGTTCAATC	ACCTGCAAGA	CGAAGAGGCA	120
AGG ATG CTG TTG GCC TGG GTA CAA GCA TTC CTC GTC AGC AAC ATG CTC	168					
Met Leu Leu Ala Trp Val Gln Ala Phe Leu Val Ser Asn Met Leu						
-21 -20 -15 -10						
CTA GCA GAA GCC TAT GGA TCT GGA GGC TGT TTC TGG GAC AAC GGC CAC	216					
Leu Ala Glu Ala Tyr Gly Ser Gly Gly Cys Phe Trp Asp Asn Gly His						
-5 1 5 10						
CTG TAC CGG GAG GAC CAG ACC TCC CCC GCG CCG GGC CTC CGC TGC CTC	264					
Leu Tyr Arg Glu Asp Gln Thr Ser Pro Ala Pro Gly Leu Arg Cys Leu						
15 20 25						
AAC TGG CTG GAC GCG CAG AGC GGG CTG GCC TCG GCC CCC GTG TCG GGG	312					
Asn Trp Leu Asp Ala Gln Ser Gly Leu Ala Ser Ala Pro Val Ser Gly						
30 35 40						
GCC GGC AAT CAC AGT TAC TGC CGA AAC CCG GAC GAG GAC CCG CGC GGG	360					
Ala Gly Asn His Ser Tyr Cys Arg Asn Pro Asp Glu Asp Pro Arg Gly						
45 50 55						
CCC TGG TGC TAC GTC AGT GGC GAG GCC GGC GTC CCT GAG AAA CGG CCT	408					
Pro Trp Cys Tyr Val Ser Gly Glu Ala Gly Val Pro Glu Lys Arg Pro						
60 65 70						
TGC GAG GAC CTG CGC TGT CCA GAG ACC ACC TCC CAG GCC CTG CCA GCC	456					
Cys Glu Asp Leu Arg Cys Pro Glu Thr Thr Ser Gln Ala Leu Pro Ala						
75 80 85 90						
TTC ACG ACA GAA ATC CAG GAA GCG TCT GAA GGG CCA GGT GCA GAT GAG	504					
Phe Thr Thr Glu Ile Gln Glu Ala Ser Glu Gly Pro Gly Ala Asp Glu						
95 100 105						
GTG CAG GTG TTC GCT CCT GCC AAC GCC CTG CCC GCT CGG AGT GAG GCG	552					
Val Gln Val Phe Ala Pro Ala Asn Ala Leu Pro Ala Arg Ser Glu Ala						
110 115 120						
GCA GCT GTG CAG CCA GTG ATT GGG ATC AGC CAG CGG GTG CGG ATG AAC	600					
Ala Ala Val Gln Pro Val Ile Gly Ile Ser Gln Arg Val Arg Met Asn						
125 130 135						
TCC AAG GAG AAA AAG GAC CTG GGA ACT CTG GGC TAC GTG CTG GGC ATT	648					
Ser Lys Glu Lys Lys Asp Leu Gly Thr Leu Gly Tyr Val Leu Gly Ile						
140 145 150						
ACC ATG ATG GTG ATC ATC ATT GCC ATC GGA GCT GGC ATC ATC TTG GGC	696					
Thr Met Met Val Ile Ile Ile Ala Ile Gly Ala Gly Ile Ile Leu Gly						
155 160 165 170						
TAC TCC TAC AAG AGG GGG AAG GAT TTG AAA GAA CAG CAT GAT CAG AAA	744					
Tyr Ser Tyr Lys Arg Gly Lys Asp Leu Lys Glu Gln His Asp Gln Lys						
175 180 185						
GTA TGT GAG AGG GAG ATG CAG CGA ATC ACT CTG CCC TTG TCT GCC TTC	792					
Val Cys Glu Arg Glu Met Gln Arg Ile Thr Leu Pro Leu Ser Ala Phe						
190 195 200						
ACC AAC CCC ACC TGT GAG ATT GTG GAT GAG AAG ACT GTC GTG GTC CAC	840					
Thr Asn Pro Thr Cys Glu Ile Val Asp Glu Lys Thr Val Val Val His						
205 210 215						

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

Sub
a1
cont.

ACC AGC CAG ACT CCA GTT GAC CCT CAG GAG GGC AGC ACC CCC CTT ATG 888
 Thr Ser Gln Thr Pro Val Asp Pro Gln Glu Gly Ser Thr Pro Leu Met
 220 225 230
 GGC CAG GCC GGG ACT CCT GGG GCC T GAGCCCCCCC AGTGGGCAGG 933
 Gly Gln Ala Gly Thr Pro Gly Ala
 235 240
 AGCCCATGCA GAACTGGTG CAGGACAGCC CACCCTCCTA CAGCTAGGAG GAACTACCAC 993
 TTTGTGTTCT GGTAAAACC CTACCACTCC CCCGCTTTTT TGGCGAATCC TAGTAAGAGT 1053
 GACAGAAGCA GGTGGCCCTG TGGGCTGAGG GTAAGGCTGG GTAGGGTCCT AACAGTGCTC 1113
 CTTGTCCATC CTTGGAGCA GATTTTGTCT GTGGATGGAG ACAGTGGCAG CTCCCACAGT 1173
 GATGCTGCTG CTAAGGGCTT CCAAACATTG CCTGCACCCC TGGAAGTGAA CCAGGGATAG 1233
 ACGGGGAGCT CCCCAGGCT CCTCTGTGCT TTAATAAGAT GGCTCAGTCT CCACTGTGGG 1293
 CTTGAGTGGC ATACACTGTT ATTCATGGTT AAGGTAAAGC AGGTCAAGGG ATGGCATTGA 1353
 AAAAATATAT TTAGTTTTTA AAATATTTGG GATGGAACTC CCTACTGACC TCTGACAACT 1413
 GGAAACGAGT TTGTACTGAA GTCAGAACTT TGGGTTGGGA ATGAGATCTA GGTGTGGCT 1473
 GCTGGTATGC TTCAGCTTGC TGGCAATGAT GTGCCTTGAC AACCGTGGGC CAGGCCTGGG 1533
 CCCAGGGACT CTTCTGTTT CATAAGGAAA GGAAGAATTG CACTGAGCAT TCCACTTAGG 1593
 AAGAGGATAG AGAAGGATCT GTCCTGCCTT TGGCCACAGG AGCAGAGGCA GACCTGGGAT 1653
 GCCCCAGTTT CTCTTCAGGG ATGGATAGTG ACCTGTCTTC ATTTTGCACA GGTAAGAGAG 1713
 TAGTTAGCTA ACCTATGGGA ATTATACTGT GGGGCTTGT GAGCTGCTTC TAAGAGGCTA 1773
 ACCTGGAAAC TAAGCTCAGA GGCAAGGTAA TAAAGCACTT CAGGGCTTGC TCCCCAAGTG 1833
 GGCCTGATTT AGCAGGTGGT CTGCGGGCGT CCAGGTCAGC ACCTTCCTGT AGGGCACTGG 1893
 GGCTAGGGTC ACAGCCCCTA ACTCATAAAG CAATCAAAGA ACCATTAGAA AGGGCTCATT 1953
 AAGCCTTTTG GACACAGGAC CCCAGAGAGG AAAAAGTGAC TTGCCCAAGG TCGTAAGCAA 2013
 GCTACTGGCA TGGCAAGAGC CCAGCTTCCT GACGGAGCGC AACATTTCTC CACTGCACTG 2073
 TGCTAGCAGC TCAGCAGGGC CTCTAACCTG TGATGTCACA CTCAAGAGGC CTTGGCAGCT 2133
 CCTAGCCATA GAGCTTCCTT TCCAGAACCC TTCCACTGCC CAATGTGGAG ACAGGGGTTA 2193
 GTGGGGCTTT CTATGGAGCC ATCTGCTTTG GGGACCTAGA CCTCAGGTGG TCTCTTGGTG 2253
 TTAGTGATGC TGGAGAAGAG AATATTACTG GTTCTACTT TTCTATAAAG GCATTTCTCT 2313
 ATAAAAAAAA AAAAAA 2329

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 263 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

Sub
a!
cont.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Leu Leu Ala Trp Val Gln Ala Phe Leu Val Ser Asn Met Leu Leu
 -21 -20 -15 -10

Ala Glu Ala Tyr Gly Ser Gly Gly Cys Phe Trp Asp Asn Gly His Leu
 -5 1 5 10

Tyr Arg Glu Asp Gln Thr Ser Pro Ala Pro Gly Leu Arg Cys Leu Asn
 15 20 25

Trp Leu Asp Ala Gln Ser Gly Leu Ala Ser Ala Pro Val Ser Gly Ala
 30 35 40

Gly Asn His Ser Tyr Cys Arg Asn Pro Asp Glu Asp Pro Arg Gly Pro
 45 50 55

Trp Cys Tyr Val Ser Gly Glu Ala Gly Val Pro Glu Lys Arg Pro Cys
 60 65 70 75

Glu Asp Leu Arg Cys Pro Glu Thr Thr Ser Gln Ala Leu Pro Ala Phe
 80 85 90

Thr Thr Glu Ile Gln Glu Ala Ser Glu Gly Pro Gly Ala Asp Glu Val
 95 100 105

Gln Val Phe Ala Pro Ala Asn Ala Leu Pro Ala Arg Ser Glu Ala Ala
 110 115 120

Ala Val Gln Pro Val Ile Gly Ile Ser Gln Arg Val Arg Met Asn Ser
 125 130 135

Lys Glu Lys Lys Asp Leu Gly Thr Leu Gly Tyr Val Leu Gly Ile Thr
 140 145 150 155

Met Met Val Ile Ile Ile Ala Ile Gly Ala Gly Ile Ile Leu Gly Tyr
 160 165 170

Ser Tyr Lys Arg Gly Lys Asp Leu Lys Glu Gln His Asp Gln Lys Val
 175 180 185

Cys Glu Arg Glu Met Gln Arg Ile Thr Leu Pro Leu Ser Ala Phe Thr
 190 195 200

Asn Pro Thr Cys Glu Ile Val Asp Glu Lys Thr Val Val Val His Thr
 205 210 215

Ser Gln Thr Pro Val Asp Pro Gln Glu Gly Ser Thr Pro Leu Met Gly
 220 225 230 235

Gln Ala Gly Thr Pro Gly Ala
 240

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 372 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

Sub
a
cont.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Tyr Val Phe Lys Ala Gly Lys Tyr Ser Ser Glu Phe Cys Ser Thr Pro
 1 5 10 15
 Ala Cys Ser Glu Gly Asn Ser Asp Cys Tyr Phe Gly Asn Gly Ser Ala
 20 25 30
 Tyr Arg Gly Thr His Ser Leu Thr Glu Ser Gly Ala Ser Cys Leu Pro
 35 40 45
 Trp Asn Ser Met Ile Leu Ile Gly Lys Val Tyr Thr Ala Gln Asn Pro
 50 55 60
 Ser Ala Gln Ala Leu Gly Leu Gly Lys His Asn Tyr Cys Arg Asn Pro
 65 70 75 80
 Asp Gly Asp Ala Lys Pro Trp Cys His Val Leu Lys Asn Arg Arg Leu
 85 90 95
 Thr Trp Glu Tyr Cys Asp Val Pro Ser Cys Ser Thr Cys Gly Leu Arg
 100 105 110
 Gln Tyr Ser Gln Pro Gln Phe Arg Ile Lys Gly Gly Leu Phe Ala Asp
 115 120 125
 Ile Ala Ser His Pro Trp Gln Ala Ala Ile Phe Ala Lys His Arg Arg
 130 135 140
 Ser Pro Gly Glu Arg Phe Leu Cys Gly Gly Ile Leu Ile Ser Ser Cys
 145 150 155 160
 Trp Ile Leu Ser Ala Ala His Cys Phe Gln Glu Arg Phe Pro Pro His
 165 170 175
 His Leu Thr Val Ile Leu Gly Arg Thr Tyr Arg Val Val Pro Gly Glu
 180 185 190
 Glu Glu Gln Lys Phe Glu Val Glu Lys Tyr Ile Val His Lys Glu Phe
 195 200 205
 Asp Asp Asp Thr Tyr Asp Asn Asp Ile Ala Leu Leu Gln Leu Lys Ser
 210 215 220
 Asp Ser Ser Arg Cys Ala Gln Glu Ser Ser Val Val Arg Thr Val Cys
 225 230 235 240
 Leu Pro Pro Ala Asp Leu Gln Leu Pro Asp Trp Thr Glu Cys Glu Leu
 245 250 255
 Ser Gly Tyr Gly Lys His Glu Ala Leu Ser Pro Phe Tyr Ser Glu Arg
 260 265 270
 Leu Lys Glu Ala His Val Arg Leu Tyr Pro Ser Ser Arg Cys Thr Ser
 275 280 285
 Gln His Leu Leu Asn Arg Thr Val Thr Asp Asn Met Leu Cys Ala Gly
 290 295 300
 Asp Thr Arg Ser Gly Gly Pro Gln Ala Asn Leu His Asp Ala Cys Gln
 305 310 315 320
 Gly Asp Ser Gly Gly Pro Leu Val Cys Leu Asn Asp Gly Arg Met Thr
 325 330 335
 Leu Val Gly Ile Ile Ser Trp Gly Leu Gly Cys Gly Gln Lys Asp Val

62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

Sub
 a.
 cont.

340 345 350
 Pro Gly Val Tyr Thr Lys Val Thr Asn Tyr Leu Asp Trp Ile Arg Asp
 355 360 365
 Asn Met Arg Pro
 370

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 250 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ATTGCACTGA GCATTCCACT TAGGAAGAGG ATAGAGAAGG ATCTGCTCCG CCTTTGGCCA	60
CAGGAGCAGA GGCAGACCTG GGATGCCCA TTTCTCTTCA GGGATGGATA GTGACCTGTC	120
TTCATTTTGC ACAGGTAAGA GAGTAGTTAG CTAACCTATG GGAATTATAC TGTGGGGCCT	180
TGTAGCTGCT TCTAAGAGGC TAACCTGGAA ACTAAGCTCA GAGGCAAGGT AATAAAGCAC	240
TTCAGGGCTT	250

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 247 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATAGAGAAAT GCCTTTATAG AAAAGTAGAA ACCAGTAATA TTCTCTTCTC CAGCATCACT	60
AACACCAAGA GACCACCTGA GGTCTAGGTC CCCAAAGCAG ATGGCTCCAT AGAAAGCCCC	120
ACTAACCCGT CTCCACATTG GGCAGTGGAA GGGTTCTGGA AAGGAAGCTC TATGGCTAGG	180
AGCTGCCAAG GCCTCTTGAG TGTGACATCA CAGGTTAGAG GCCCTGCTGA GCTGCTAGCA	240
CAGTGCA	247

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 461 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

340 345 350 355 360 365 370

Sub
 a!
 Cont.

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AATTCGGCAA GAGTAACAGC ATAACAAGGG TAGGTCTGAC TGCAGCTGGG ACTGGGAGGC	60
AGAGCACGCC AAGGGGGCCT CGGTAAACA CTGGTCGTTT AATCACCTGC AAACGAGGAG	120
GCAAGGATGC TGTTGGCCTG GGTACAGCAT TCCTGGTCAG CAACATGCTC CTAGCGTAAG	180
CCTATGGATC TGGAGGCTGT TTCTGGGACA ACGGCCACTG TACCCGGAGG ACCAGACCTT	240
CCCGGCCGGT CCTCGTGCCT CAACTGGCTG GACGCGCAGG GCTGCCTGGG CCCCTTTTC	300
GGTCAAATTT CACAGTTTAC TTCGAAACCG GGACGGGGCC GTGGGGGCC TGGTGGTTAG	360
TTTGGGGTCG GGTTTTCTTA AAAAAGGTTT TTGGGGCCGG TTTTCGGAAC CATTTTCGGT	420
GAATTTTTTA GGGAAATTTC AGGAGTTTTT TAAGGGCCAT T	461

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 399 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGCAAGTTGC AGAACTGGAA ACGAGTTTGT ACAGAAGTCA GAACTTTGGG TTAGGAATGA	60
GATCTAGGTT GTGGCTGCTG GTATGCTTCA TTGCTGGCAA TAATGTGCCT TGACAACCGT	120
GGGCCAGGCC TGGGACCAGG GACTCTTCCT GTTTCATAAG GAAAGGAAGA ATTGCACTGA	180
GCATTCCACT TAGGAAGAGG ATAGAGCAAG GAATCTGCTC CGCTTTGGCC ACAGGAGCAG	240
AGGCAGACCT GGGATGCCCC AGTTCTCTTT CAGGGATGGG ATAGTGACCT GTCTTACATT	300
TTGCACAGGT AAAGAGAGTT AGTTAGCTAA CCTATTGGGC TTTATTACTT GGGGCTTGTC	360
AGCTGCTTTT TAAGAGGTTA ACCTGGAACCT AAAGTTCAG	399

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 334 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

66443000

*Sub
a
cont.*

TAATTCGGCA AGGGACAGGT CTGACTGCAG CTGGGACTGG GAGGCAGAGC CGTCAAGGGG 60
 GCCTCGGTTA AACACTGGTC GTTCAATCAC CTGCAACGAG AGGCAAGGAT GCTGTTGGCC 120
 TGGGTACAAG CATTCTGTGC AGCAACATGC TCCTAGCAGA AAGCCTATGG ATCTGGGAGG 180
 CTGTTTCTGG GACAACGGCC ACCTGTACCG GAGGACCAGA CCTCCCCGGC CGGGCCTTCC 240
 GTGGCCTTCA ATTGGTTTGA CGTGGCAAAG GGGCTTGTCT GGCCCTTTTG GGGGAAAATT 300
 ACAAGTTTTA ATTGTCCCGG AAAACCTGGA GAGG 334

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 472 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AATTCGGCAG AGGGAGAGGG AGATGCAGCG AATCACTCTG CCCTTGTCTG CCTTCACCAA 60
 CCCACCTGT GAGATTGTGG ATGAGAAGAC TGTCGTGGTC CACACCAGCC AGACTCCAGT 120
 TGACCCTCAG GAGGGCAGCA CCCCCCTTAT GGGACCAGGC CGGGGACTCC TGGGGCCTGA 180
 GCCCCCAGT GGGGCAGGAG CCATGGCAGA CACTGGTGCA GGACAGCCAC CCTCCTTACA 240
 GCTAGGGGGA ACTACCACTT TGTGTTTCTG GTTTAAACC CTACCACTCC CGGATTTTTT 300
 GCGGATTCC TTAGTTAAGA GTACAGAAGC AGGTGGGCCT ATGGCTTGGA GGGTAAGGTG 360
 GGGTAGGGTT CCTAAAAGTG GGTTCCTGGT TGCTCCTGGG AGGAAGATTT TGGTTTGGT 420
 GGGGACAGTG GCAGTTTCCA CAGGTTGTTG TGTTAAGGGG TTCAAAAAAT TG 472

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 291 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGGCACGAGA TGAACCTCAA GGAGAAAAG GACCTGGGAA CTCTGGGTAT GACGGTCCCC 60
 CACCCCTGCC CTTGTTGGGA TTCATCAAGA GATGTCATTT GCTGATTGTC TAGGGTGTGG 120
 CTAATGGGAC CTTGTGTCCT ATCCTTGGCA GGCTACGTGC TGGGCATTAC CATGATGGTG 180
 ATCATCATTG CCATCGGAGC TGGCATCATC TTGGGCTACT CTACAAGAGG TCAGTAGCTT 240

*Sub
a1
cont*

CTCTTCTGGG CCTCTTAGG AGGAGGGGAG GAAGGTACAC AAAGTCAAAC T

291

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGCCGACATG TCTGGAGGCT GTTTCTGG

28

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GGCGGAAGCT TATTAGGCC CAGGAGTCCC GGC

33

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GGCCGGGATC CGCCATCATG CTGTTGGCCT GGGTAC

36

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 35 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

GGCCGGGATC CGCCATCATG CTGTTGGCCT GGGTAC

Sub
a1
cont.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GGCCGGGTAC CTTATTAGGC CCCAGGAGTC CCGGC

35

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GGCCGGGATC CGCCATCATG CTGTTGGCCT GGGTAC

36

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GGCCGGGTAC CTTATTAGGC CCCAGGAGTC CCGGC

35

Sub
al
cont